OPEN ACCESS
Saudi Journal of Engineering and Technology
Abbreviated Key Title: Saudi J Eng Technol
ISSN 2415-6272 (Print) |ISSN 2415-6264 (Online)
Scholars Middle East Publishers, Dubai, United Arab Emirates
Journal homepage: https://saudijournals.com

Original Research Article

Development of a Hybrid Support Vector Machine and Bat Optimization Algorithm for Enhanced Spectrum Sensing

Abubakar Yahaya Hamisu[1]*, Matthew Ehikhamenle[1]

[1]Center for Information and Telecommunication Engineering, University of Port Harcourt, School of Graduate Studies, Rivers State, Nigeria

DOI: https://doi.org/10.36348/sjet.2026.v11i09.005 | **Received:** 03.07.2026 | **Accepted:** 27.08.2026 | **Published:** 14.09.2026

***Corresponding author:** Abubakar Yahaya Hamisu
Center for Information and Telecommunication Engineering, University of Port Harcourt, School of Graduate Studies, Rivers State, Nigeria

Abstract

This paper presents the findings of a research on an artificial intelligence-based hybrid Support Vector Machine–Bat Optimization Algorithm (SVM-BOA) model developed for spectrum sensing in the Ultra High Frequency (UHF) television band (470 – 694 MHz). The main objective is to enhance detection accuracy, probability of detection, probability of false alarm, and computational efficiency for efficient spectrum utilization. Experimental spectrum measurements were conducted across Kano, Kaduna, and the Federal Capital Territory (FCT), Abuja, Nigeria, to assess the level of spectrum occupancy and identify available Television White Space (TVWS) channels. Relevant signal features, including received signal power level, noise power level, and frequency, were extracted from the field data collected and used to train and test the proposed model. The model was developed and implemented in MATLAB computing environment. The new model was validated using test data for classification of channel status as either free or occupied. The developed model achieved a detection accuracy of 97.67% at a threshold of −99.5dB, based on metric from confusion matrix, outperforming baseline SVM (78.57%) and other benchmark optimization-based SVM models, including Grid Search SVM (92.86%), Random Search SVM (89.29%), Genetic Algorithm SVM (92.86%), and Particle Swarm Optimization SVM (92.78%). It also attained a probability of detection (Pd) of 1.00 (100%) and a probability of false alarm (Pfa) of 0.1053 (10.53%), which satisfies the sensing performance requirements of the Institute of Electrical and Electronics (IEEE 802.22) standards under low-signal conditions. These findings suggest that the hybrid SVM-BOA approach will offer superior classification accuracy and computational efficiency relative to conventional techniques and comparable hybrid models, signifying a strong potential for the improvement of spectrum utilization in next-generation cognitive radio networks.

Keywords: Television White Spaces (TVWS); Spectrum Measurement; Spectrum Sensing; Support Vector Machine (SVM); Bat Algorithm; Cognitive Radio.

1. INTRODUCTION

The exponential growth of wireless applications and data-centric services has placed unprecedented demand on the radio frequency spectrum, a finite and inherently scarce natural resource. The traditional command-and-control model of spectrum allocation, while providing interference-free communication, has led to a paradox: while the spectrum is perceived as congested, significant portions remain vastly underutilized. This is particularly acute in the Ultra High Frequency (UHF) bands (470–694 MHz), originally allocated for Television Broadcasting. The underutilization of these prime frequencies, characterized by superior propagation characteristics, represents a critical inefficiency within the existing spectrum management framework. This condition is particularly pronounced within the Nigerian telecommunication landscape. Investigation on spectrum occupancy has indicated that a significant portion of the UHF spectrum particularly within the Sub-urban and Rural Areas remains underutilized as Television White Space (TVWS), Ubom, & Ukommi, (2023). A study on spectrum occupancy measurement and analysis in Ogun State Nigeria by Lasisi *et al.,* (2022) revealed that approximately 60% of the UHF spectrum in urban areas and nearly 100% in rural areas lies underutilized as

White Space. This vast availability of unused spectrum underscores the urgent need for a paradigm shift from static allocation to Dynamic Spectrum Access (DSA), enabled by robust Dynamic Spectrum Sensing (DSS) Techniques. The fundamental objective of DSS is to reliably detect the presence or absence of Primary Users (PUs), potentially facilitating opportunistic access for Secondary Users (SUs) without inducing harmful interference.

However, the development of an effective DSS framework for TVWS in such real-world environments is confronted by several critical and interconnected challenges: (i) Vulnerability to Noise Uncertainty and Low SNR, (ii) Inadequacy of Simple Error Detection Models. In spectrum sensing, detection performance is characterized by two primary error types: Missed Detection and False Alarm. These are modeled using the following probabilities: Probability of Detection (P_d): The probability that the system correctly identifies the presence of a PU signal when the spectrum is occupied (H_1). Probability of False Alarm (P_{fa}): The probability that the system incorrectly identifies a PU signal when the spectrum is actually idle (H_0) (Yang, X.-S., He, X. 2023). These models must account for the trade-off between the probability of false alarm (P_{fa}), which result in spectrum underutilization by incorrectly identifying a vacant channel as occupied and the probability of missed detection (P_{md}), which could result to harmful interference by failing to detect an active PU. The primary objective is to minimize both errors, however, traditional methods often fail to maintain this balance under dynamic channel conditions, resulting in either inefficient use of the TVWS or unacceptable levels of interference to incumbent television broadcasters.

To overcome the limitations of traditional methods, advanced Machine Learning (ML) and Artificial Intelligence (AI) models, such as Support Vector Machines (SVM), has been proposed. While SVMs can offer superior detection accuracy, especially in non-Gaussian noise, their performance is highly sensitive to the choice of hyper-parameters and the selection of relevant features from the raw signal data. Consequently, the optimization of these parameters is often computationally expensive. This dynamic suggests a significant computational cost trade-off: highly accurate models may prove unsuitable for real-time sensing applications wherein speed and power efficiency are paramount (Wainer & Fonseca2021). Therefore, there is a clear need for an intelligent optimization mechanism that can automatically tune these models for high accuracy without incurring prohibitive computational overhead. (iv) Lack of Context-Specific Data and Frameworks in Nigeria: The performance of any data-driven model is intrinsically linked to the data on which it is trained. A major impediment to the development of a robust DSS solutions for Nigeria is the limitation of available datasets. Most existing models are trained on synthetically generated signals or measurements from temperate regions, which do not accurately reflect the unique signal propagation characteristics, noise floor dynamics, and interference patterns prevalent in a tropical country like Nigeria. Recent studies suggest that the core problem is the absence of an efficient framework for DSS in TVWS, which is further compounded by the lack of models validated with real-world spectrum measurement data from the Nigerian environment, Lasisi etal., (2022) This data gap renders many existing solutions potentially ineffective when deployed locally.

Recent literatures revealed that, Support Vector Machine (SVM) models facilitate the construction of an optimal hyperplane, that effectively separate distinct classes within the feature space; the decision function is generally expressed mathematically as presented in equation (1), (Yang, X.-S., He, X. 2023)

$$f\ (x) = sgn\ \left(\sum_{i=1}^{n} \alpha_i y_i K(x,\ x_i) + b\right) \qquad (1)$$

Where K (x, x_i) is the kernel function that facilitate quantification of the similarity between input data points.

The Bat Algorithm (BA) is used to optimize the SVM parameters (specifically the penalty factor C and the kernel parameter γ), which directly shifts the decision boundary to balance the trade-off between the probability of detection (P_d) and the probability of false alarm (P_{fa}) (Yang, X.-S., He, X. 2023). An optimized hybrid model seeks to maximize P_d (to protect the PU) while maintaining P_{fa} below the regulatory threshold typically established at 0.1 according to IEEE 802.22 standards. (iii) Computational Cost vs. Real-Time Performance Trade-offs:

The fundamental challenge therefore, is the lack of an efficient, accurate, and context-aware framework for dynamic spectrum sensing in TVWS networks. Existing methods are hampered by the combined challenges of noise uncertainty, inadequate error management, high computational costs, and a critical dependence on non-representative training data and limited scholarly study regarding hybrid optimization algorithms tailored towards spectrum sensing (Zeng, Y. Liang, Y. Hoang, A.T. & Zhang, R., 2010). To address these multifaceted challenges, the present study proposes the development of a novel hybrid model that integrates a Support Vector Machine (SVM) with a Bat Optimization Algorithm. This hybridization aims to intelligently optimize the SVM's parameters, thereby enhancing detection accuracy in low SNR conditions, managing detection errors effectively, and achieving a favourable trade-off in computational cost. The proposed model will be developed and validated using real-world spectrum measurement data from Nigeria, ensuring its practical applicability and contribution to the efficient utilization of the nation's scarce spectrum resources.

2. METHODOLOGY: MATERIAL AND METHOD

2.1 Conceptual Framework: The Proposed Hybrid Model

The conceptual framework (Figure 1.) shows the interaction of the key components. The model is structured as a closed-loop optimization and classification system.

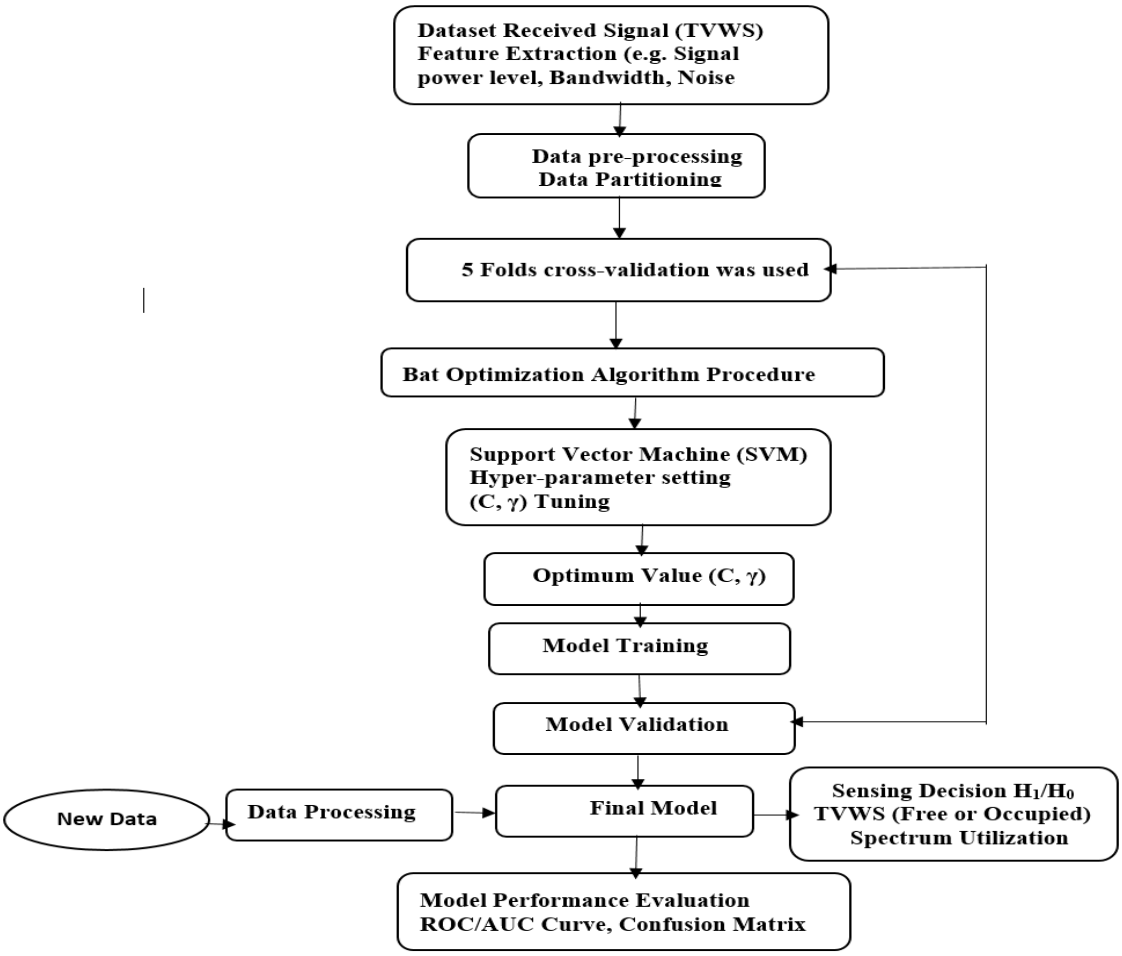


Figure 1: Block diagram of Conceptual Framework for SVM - BOA Hybrid Spectrum Sensing

The methodology comprises a five-stage process: data acquisition, feature extraction, the design and development of a hybrid Support Vector Machine-Bat Optimization Algorithm (SVM-BOA) model, training and testing, and performance evaluation (Bishnu & Bhatia, 2019; Yang & He, 2023).

I. **Data collection:** Real world TVWS data was collected via measurement campaign in three locations, namely: Kano, Kaduna and Abuja.
II. **Feature Extraction:** The extraction of salient features, comprising frequency, signal power level, noise power level, bandwidth, and occupancy status, was performed subsequent to the acquisition of field data.
III. **Total Sample Population and Data Split: Training/Validation:**
IV. A total of Eighty-four (84) samples were collected. The model was trained and validated using 5-fold cross-validation, where 4 folds (80%) were used for training and 1 fold (20%) for validation in each iteration.
V. **Signal Model Representation:** The received signal y (n) was modelled under two hypotheses (Musumeci etal., 2018):
   1. **The Null Hypothesis:** H_0: y (n) = w (n),
   2. **The Alternative Hypothesis:** H_1: y (n) = s (n) + w (n)

   Where: s (n) is the TV signal and w (n) is AWGN or impulsive noise.
VI. **The Hybrid SVM-BOA Framework**: The objective of SVM is to find the optimal hyper-plane that maximizes the margin $\frac{2}{\|w\|}$. While the Bat Optimization Loop is to (Rebentrost *et al.,* 2014):
   a. **Initialize** bat population (positions = SVM [C, γ] pairs).
   b. **Evaluate Fitness:** Fitness = Maximize P_d for a fixed P_{fa}.
   c. **Update Velocity and Frequency:** Based on echolocation behaviour.

d. **Pulse Emission & Loudness:** Adjust to switch from global search to local refinement.
e. **Termination:** Stop when the optimal values of (C, γ) are attained.

VII. **Performance Metrics:** The following metrics were used to evaluate the performance of the model:

1. **ROC/AUC Curves:** Probability of Detection (P_d) vs. False Alarm (P_{fa}).
2. **Sensing Latency:** Time taken to reach a decision.
3. **Comparison:** The methodology was compared against Support Vector Machine-Grid Search (SVM-GS), Support Vector Machine-Random Search (SVM-RS), Support Vector Machine-Genetic Algorithm (SVM-GA), Support Vector Machine-Particle Swarm Optimization (SVM-PSO), and the baseline Support Vector Machine (SVM); this comparative analysis may elucidate performance distinctions.

VIII. **Simulation Environment:** MATLAB computational environment was used as simulation tools.

IX. **Expected Results**

a. Identification of the available TVWS in the UHF TV band (470–694 MHz) across Kano, Kaduna, and Abuja (FCT) for the use of cognitive radio.
b. Design and implement a hybrid spectrum sensing algorithm integrating Support Vector Machine (SVM) with Bat Optimization Algorithm (BOA) using signal power levels measurement from the field.

2.2 Materials/Instrumentation

2.2.1 Spectrum Measurement Campaign/Data collection

The following instruments were used for the measurement campaign and data collection:

a. **Spectrum Analyzer:** Agilent (Now Keysight Technologies) Handheld Spectrum Analyzer Model No. N9344C, with a frequency range of 9 KHz to 20 GHz. This device's portability and wide frequency range made it suitable for field measurements.
b. **Antenna**: A wide band high gain (12dB) log-periodic antenna with a frequency range of 1 MHz to 26.5 GHz. This antenna was chosen for its wideband characteristics, which allowed it to cover the entire UHF TV band without requiring multiple antennas.
c. **Cables and Accessories**: Low-loss coaxial cables and a tripod for stable antenna placement were used.
d. **Amplitude Accuracy:** ± 1.3dB
e. **Frequency Range:** 9 KHz to 20 GHz (fully covering the band of interest).
f. **Frequency Accuracy:** ± (0.22 * Span + Span/ (Sweep point – 1)
g. **Frequency Span:** 100Hz – 20GHz
h. **Frequency Resolution:** 1Hz
i. **Calibration Status**: The Spectrum Analyzer was calibrated to ISO/IEC 17025:2017 standards.
j. **Key Settings:** Resolution Bandwidth (RBW) = 100 kHz, Video Bandwidth (VBW) = 30 kHz, Sweep Time = 20 Ms.
k. **Software**: Data was logged using the instrument's internal logging and later analyzed using PC-based software (Agilent BTA/DSA Software).
l. **GPS Map**: Camera Lite Application, Version 1.5.8 was adopted.

2.2.2 Measurement Locations and Site Selection

Fixed location measurements was conducted in three cities. The measurements were taken at different time of the day to capture the worst-case scenario of spectrum usage. Measurement locations were strategically chosen to represent a mix of urban, suburban, and rural environments within Kano, Kaduna, and FCT, Abuja as presented below: While, the map of figure 2 below depicts the locations of measurements.

1. **Kano:** Central City Area, (Lat: 11.98200, Long: 8.51470) representing Urban area, Kura L.G.A (Lat: 11.76050, Long: 8.4226^0), representing Sub-urban area and Gwarmai (Lat: 11.5299^0, Long: 8.2061^0), representing Rural area.
2. **Kaduna:** Central City Area (Lat: 10.52410, Long: 7.44210) representing Urban area, Zaria L.G.A (Lat: 11.1014^0, Long: 7.69300), representing Sub-urban area and Ikara (Lat: 11.41230, Long: 8.2573), representing Rural area.
3. **Abuja (FCT):** Central Business District (Lat: 9.11270, Long: 7.47690), representing Urban area, Bwari L.G.A (Lat: 9.27600, Long: 7.38540) representing Sub-urban area and Kuje (Lat: 8.45600, Long: 7.08400)**,** representing Rural area.

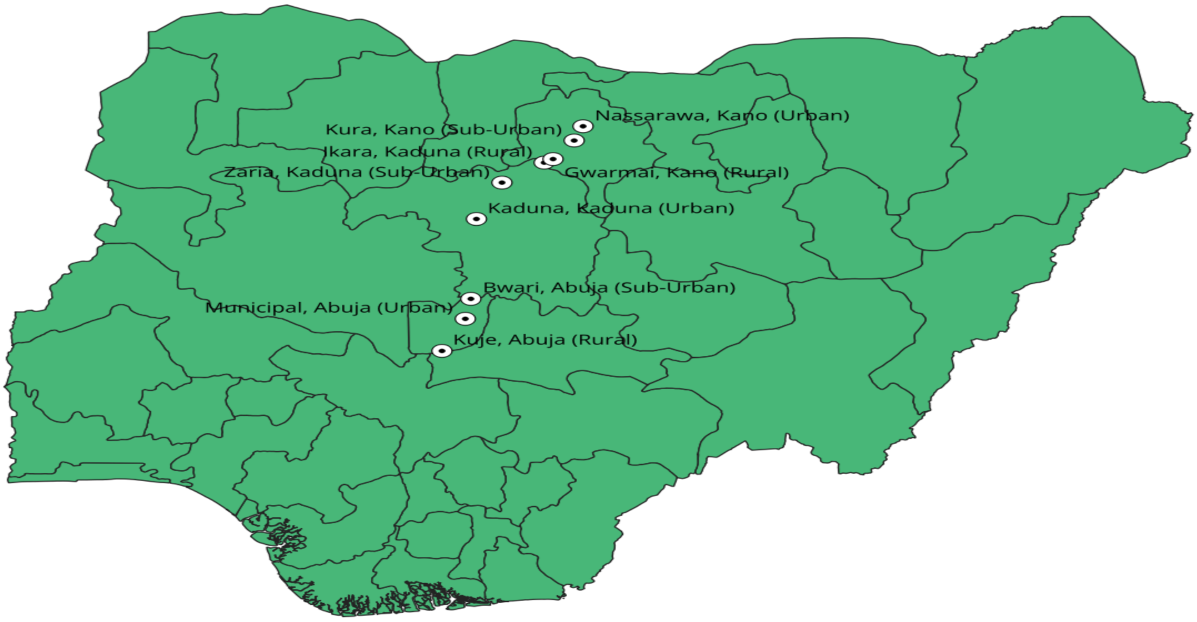


Figure 2: Map of Nigeria showing measurement locations in Kano, Kaduna and Abuja

2.2.3 Measurement Procedure

The measurement campaign followed a systematic procedure to ensure data accuracy and consistency. The spectrum analyser was connected to the log-periodic antenna using coaxial cables, and was mounted at a height of approximately 3.0 meters to minimize ground reflections and interference. Resolution Bandwidth (RBW) was set at 30 kHz to resolve individual TV channels, while the Video Bandwidth (VBW) was set at 300 kHz to smooth the display trace. Attenuation loss was set at 20dB to reduce the amplitude of the incoming signal so as to protect the instrument sensitivity front-end components like (mixers) from damage and to prevent signal distortion. The Detection Mode was set at Max Hold to capture the peak signal levels over a period. Finally, the spectrum analyser was configured to sweep the UHF television band from (470 – 694) MHz as shown in figure 3 and figure 4 below. Measurements were taken at each location for a duration of 60 minutes to capture potential intermittent transmissions. The measurements were repeated for three days and average values of the signal was calculated and recorded. The resulting spectrum traces, showing signal strength (in dBm) versus frequency, were saved for subsequent analysis.

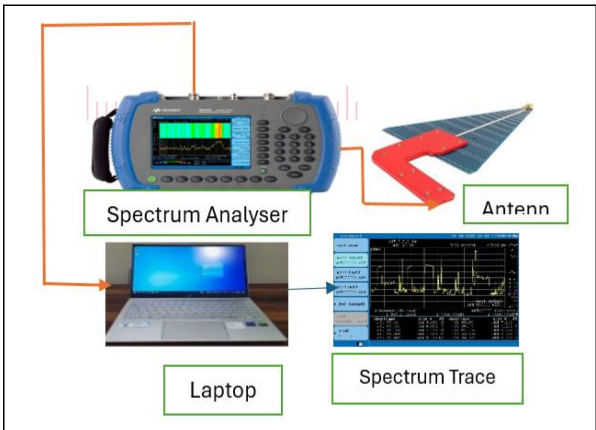


Figure 3: Spectrum Analyser connection diagram

Figure 4: Showing the Spectrum Analyser measurement set-up

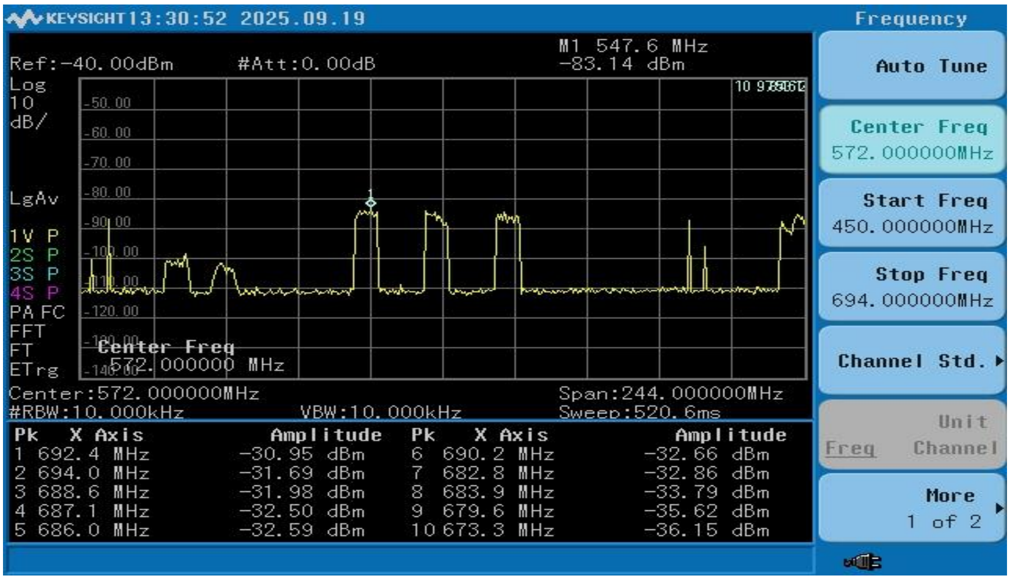
2.2.4 Data Collection and cleaning

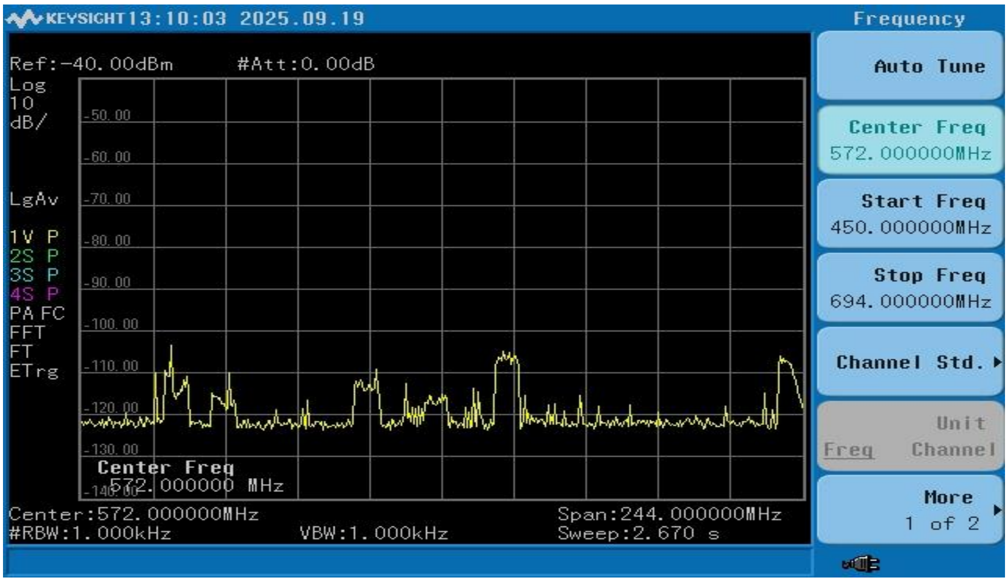
This aspect involves collection of real-world spectrum occupancy measurements data directly from primary source using the Antenna and the Spectrum Analyser as the instrument for data capture. The data was collected from the field in Kano, Kaduna and Abuja within the UHF TV band (470-694 MHz). The parameters of the data collected includes Received Signal Power Level (RSSI), frequency (Start, Centre, and End), Noise floor power level, Bandwidth, and Occupancy time. This was followed by data cleaning and preparation to ensure data quality for machine learning applications. The empirical data collected will be used for SVM – BOA model design and development. The data was categorized and class label assigned. The occupied channels were labelled as (occupied = +1) and the free channels were labelled as (free = - 1) Table 1 below present the real-world field data collected with channel and class label. Table 1 below present the real-world field data collected with channel and class label. While figures 5, 6 and 7 presents' spectrum analyser traces from Kano, Kaduna and Abuja respectively.

Table 1: Data collected/Categorization of channel and Class

Channel No.	Start Freq (MHz)	Centre Freq (MHz)	End Freq (MHz)	Average Signal power Level (dB)	Channel Labelling	Class Label
Field Data Collected from Kano						
21	470	474	478	-73.25	Occupied active	+1
22	478	482	486	-111.01	Free	-1
23	486	490	494	-75.21	Occupied active	+1
24	494	498	502	-112.26	Free	-1
25	502	506	510	-113.08	Free	-1
26	510	514	518	-100.93	Free	-1
27	518	522	526	-102.50	Free	-1
28	526	530	534	-103.75	Free	-1
29	534	538	542	-66.98	Occupied active	+1
30	542	546	550	-107.51	Free	-1
31	550	554	558	-109.04	Free	-1
32	558	562	566	-73.67	Occupied active	+1
33	566	570	574	-108.54	Free	-1
34	574	578	582	-102.54	Free	-1
35	582	586	590	-62.85	Occupied active	+1
36	590	594	598	-105.61	Free	-1
37	598	602	606	-110.20	Free	-1
38	606	610	614	-117.30	Free	-1
39	614	618	622	-111.20	Free	-1
40	622	626	630	-114.20	Free	-1
41	630	634	638	-115.30	Free	-1
42	638	642	646	-114.00	Free	-1
43	646	650	654	-115.01	Free	-1
44	654	658	662	-110.02	Free	-1
45	662	666	670	-113.20	Free	-1
46	670	674	678	-115.30	Free	-1
47	678	682	686	-66.35	Occupied active	1
48	686	690	694	-115.20	Free	-1
Field Data Collected from Kaduna						
21	470	474	478	-73.25	Occupied active	+1
22	478	482	486	-66.52	Occupied active	+1
23	486	490	494	-75.21	Occupied active	+1
24	494	498	502	-114.26	Free	-1
25	502	506	510	-77.34	Occupied active	+1
26	510	514	518	-100.93	Free	-1
27	518	522	526	-68.14	Occupied active	+1
28	526	530	534	-103.75	Free	-1
29	534	538	542	-110.23	Free	-1
30	542	546	550	-107.51	Free	-1
31	550	554	558	-109.04	Free	-1
32	558	562	566	-117.10	Free	-1
33	566	570	574	-108.54	Free	-1
34	574	578	582	-102.54	Free	-1
35	582	586	590	-114.10	Free	-1
36	590	594	598	-105.61	Free	-1
37	598	602	606	-110.20	Free	-1
38	606	610	614	-111.30	Free	-1
39	614	618	622	-111.20	Free	-1
40	622	626	630	-114.20	Free	-1

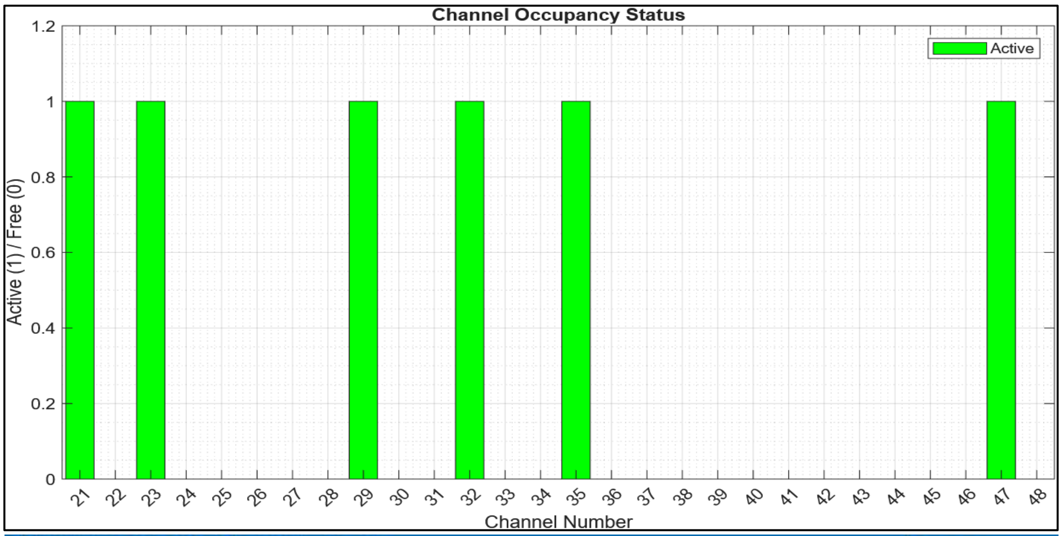
Channel No.	Start Freq (MHz)	Centre Freq (MHz)	End Freq (MHz)	Average Signal power Level (dB)	Channel Labelling	Class Label
41	630	634	638	-115.30	Free	-1
42	638	642	646	-114.00	Free	-1
43	646	650	654	-115.01	Free	-1
44	654	658	662	-110.02	Free	-1
45	662	666	670	-113.20	Free	-1
46	670	674	678	-115.30	Free	-1
47	678	682	686	-115.20	Free	-1
48	686	690	694	-83.20	Occupied active	+1
Field Data Collected from Abuja						
21	470	474	478	-85.2	Occupied active	+1
22	478	482	486	-77.34	Occupied active	+1
23	486	490	494	-84.31	Occupied active	+1
24	494	498	502	-110.73	Free	-1
25	502	506	510	-101.22	Free	-1
26	510	514	518	-109.25	Free	-1
27	518	522	526	-111.21	Free	-1
28	526	530	534	-112.22	Free	-1
29	534	538	542	-101.32	Free	-1
30	542	546	550	-118.23	Free	-1
31	550	554	558	-98.10	Occupied active	+1
32	558	562	566	-97.20	Occupied active	+1
33	566	570	574	-110.31	Free	-1
34	574	578	582	-114.51	Free	-1
35	582	586	590	-84.31	Occupied active	+1
36	590	594	598	-110.21	Free	-1
37	598	602	606	-113.22	Free	-1
38	606	610	614	-101.53	Free	-1
39	614	618	622	-113.23	Free	-1
40	622	626	630	-112.32	Free	-1
41	630	634	638	-87.10	Occupied active	+1
42	638	642	646	-100.10	Free	-1
43	646	650	654	-111.32	Free	-1
44	654	658	662	-101.54	Free	-1
45	662	666	670	-113.11	Free	-1
46	670	674	678	-112.22	Free	-1
47	678	682	686	-101.32	Free	-1
48	686	690	694	-111.31	Free	-1

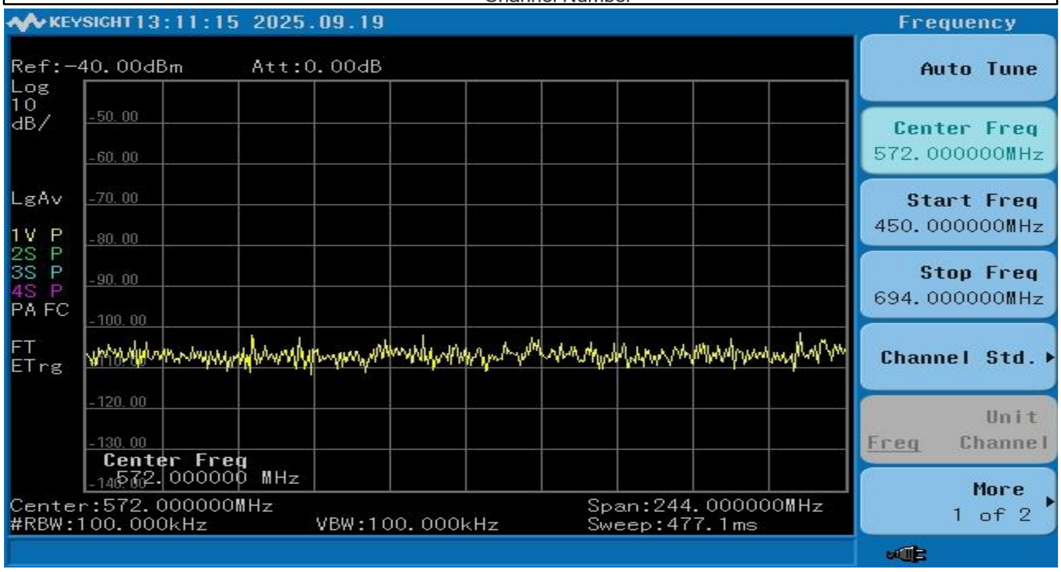




Kano Urban Area, Nasarawa GRA

Kano Sub-Urban Area, Kura

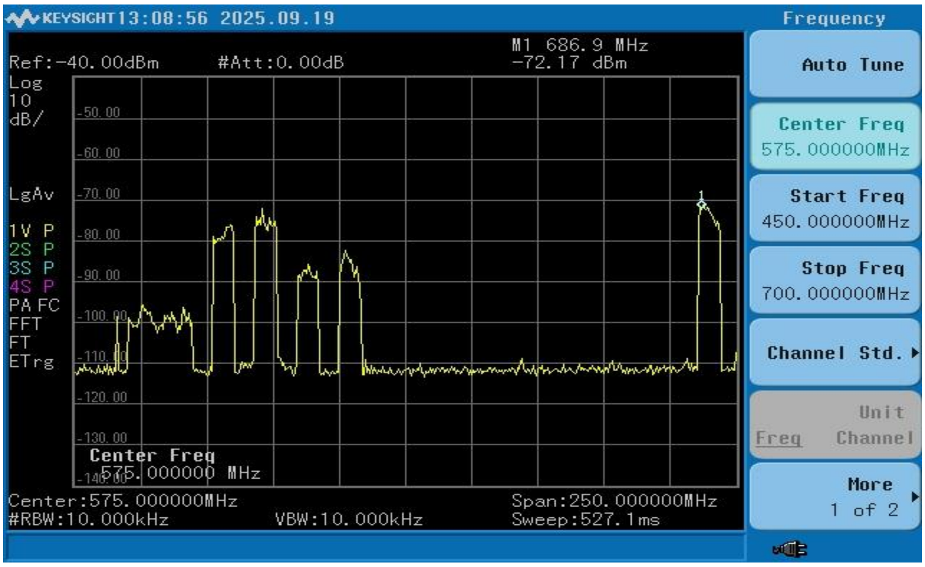




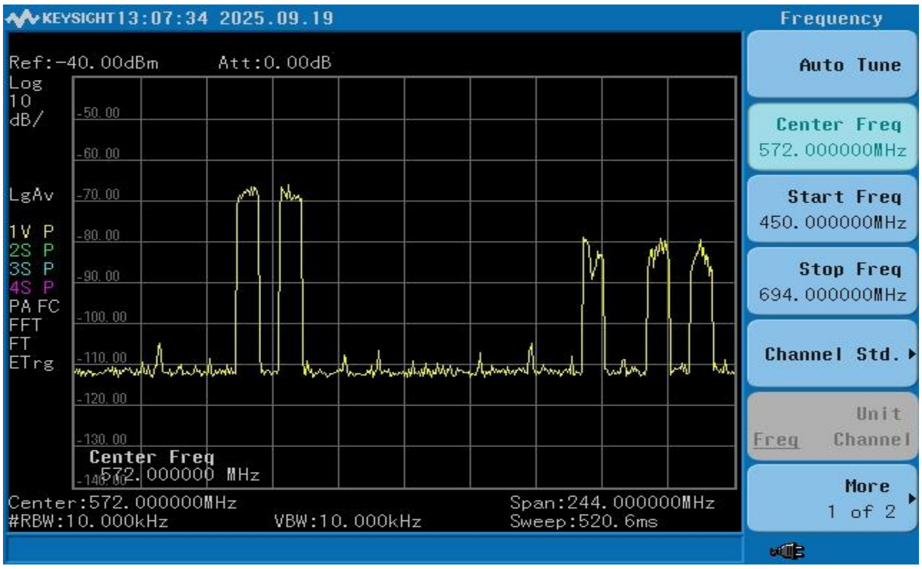
Plot of Channel against Channel Number

Kano Rural Area, Gwarmai

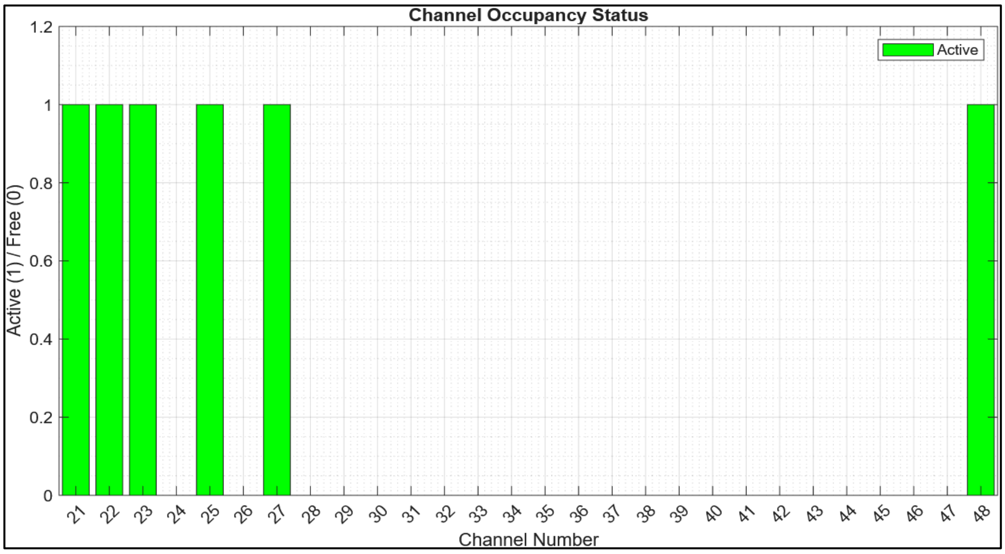
Figure 5: Spectrum Analyzer Traces from Kano



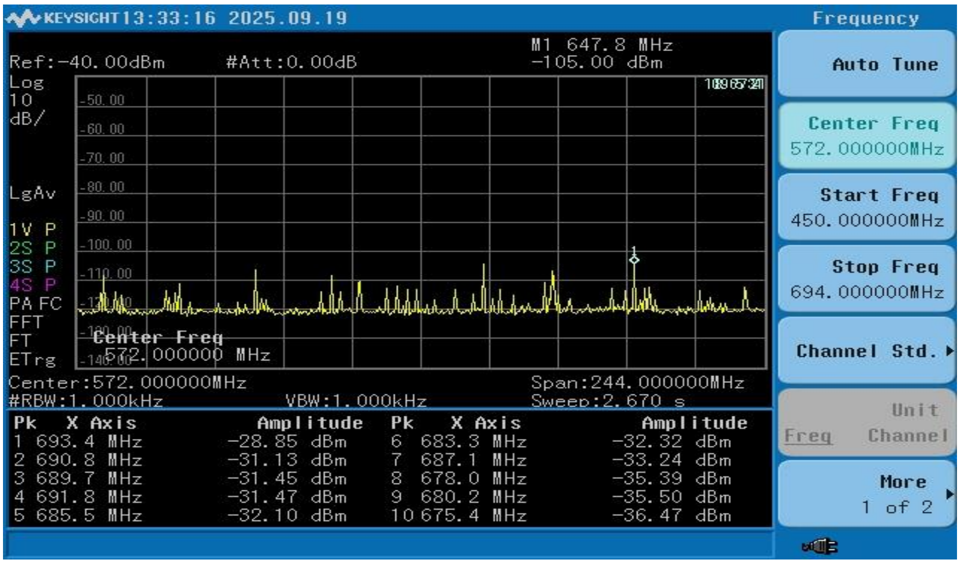
Kaduna Urban Area



Kaduna Sub-Urban Area, Zaria

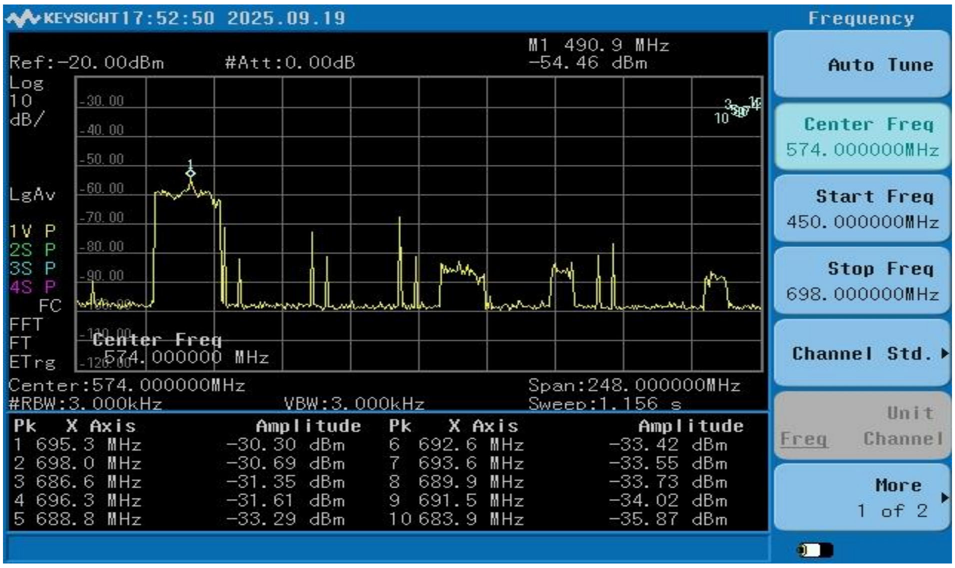


Plot of Channel against Channel Number

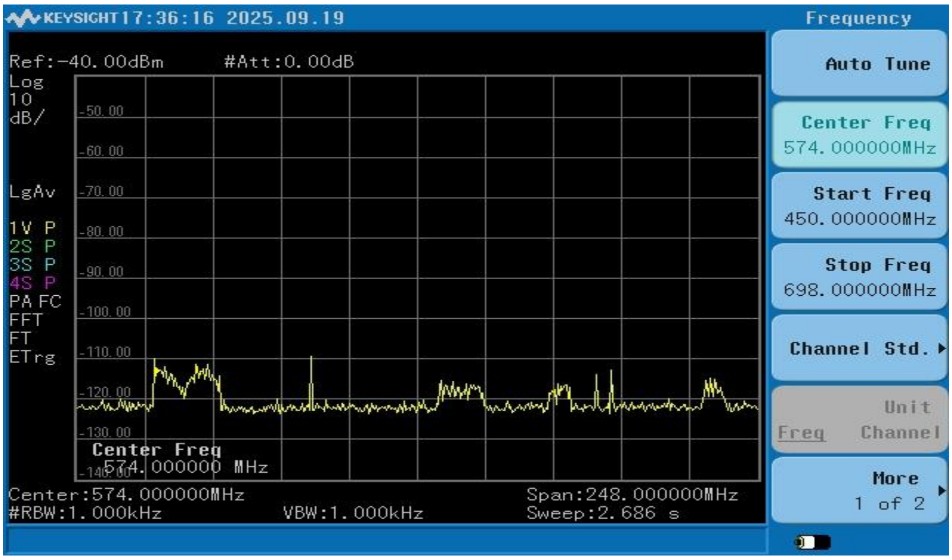


Kaduna Rural Area, Makarfi

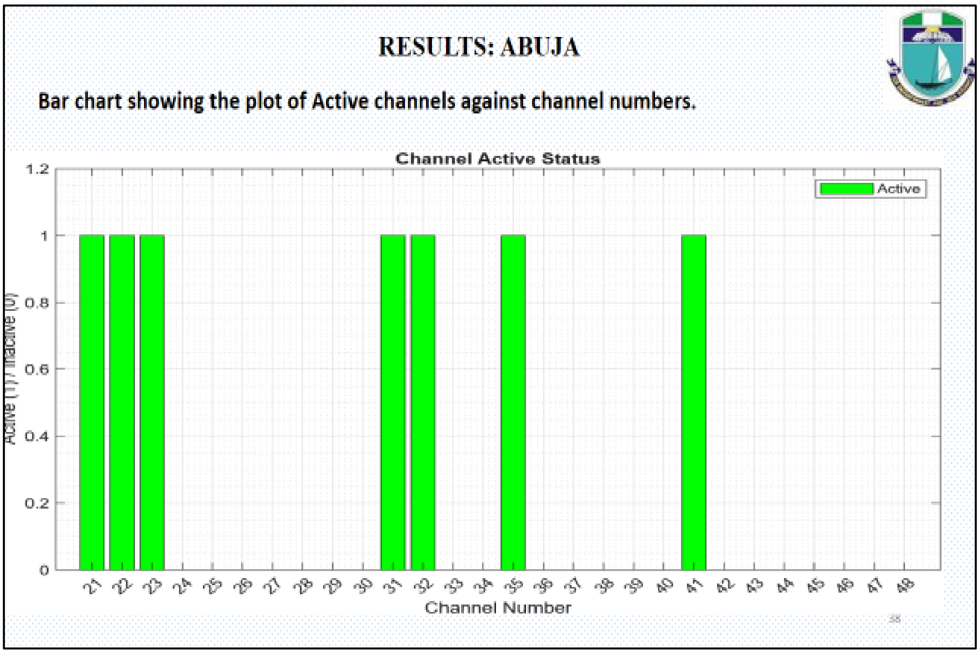
Figure 6: Spectrum Analyzer Traces from Kaduna



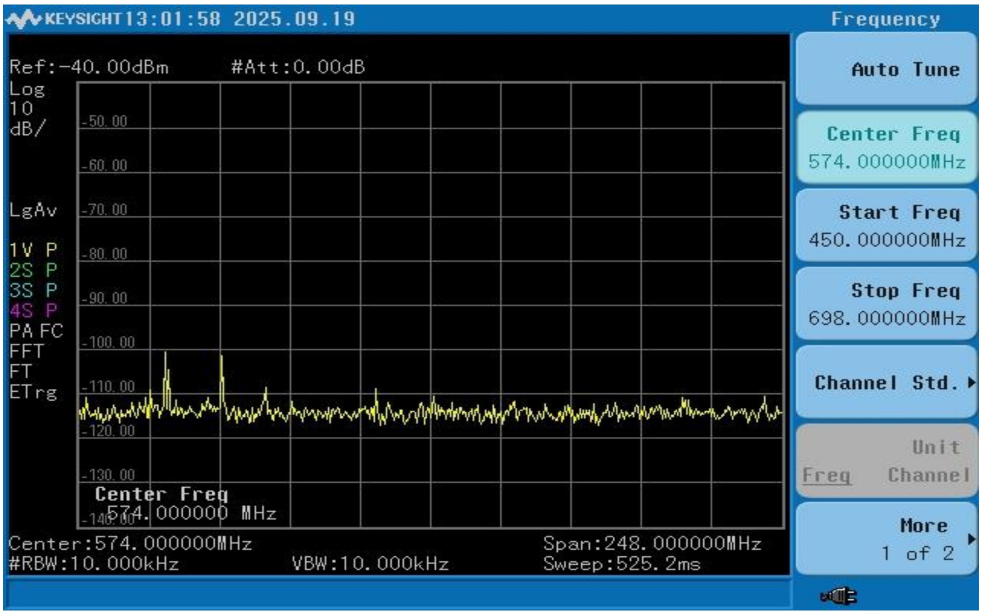
Abuja Urban Area,



Abuja Sub-Urban Area, Kuje



Plot of Channel against Channel Number



Abuja Rural Area, Buari

Figure 7: Spectrum Analyzer Traces from Abuja

2.2.5 Theoretical Framework of Cognitive Radio and TVWS

The framework of Cognitive Radio (CR), as theorized by (Haykin, S2005), revolves around the "cognitive cycle," which includes spectrum sensing, spectrum management, spectrum mobility, and spectrum sharing. Within the context of Television White Space (TVWS), spectrum sensing is subjected to strict regulatory and physical constraints. The UHF TVWS bands typically spanned (470 – 694) MHz. This range of frequency exhibit superior propagation characteristics, including low free-space path loss and high penetration capabilities through obstacles (Kaur etal., 2025). Thus, the primary users (PUs) in this band mainly Terrestrial Digital Video Broadcasting (T - DVB) or Advanced Television Systems Committee (ATSC) transmitters, together with low power wireless microphones, operate over large geographical areas.

The main mathematical objective of spectrum sensing is to choose between two competing hypotheses regarding a specific frequency band over a continuous time interval (Martin, R. Davies, A. & Collins, T., 2016)

y(t) = (t) : H_0 (spectrum free) (2)

y(t) = h(t)s(t)+ $n(t)$: H_1 (spectrum Occupied) (3)

Where:

- y(t) is the complex signal received by the Secondary User (SU).

- n(t) is Additive White Gaussian Noise (AWGN) modeled with zero mean and variance σ_n^2
- s(t) represents the transmitted PU signal with variance σ_s^2
- h(t) is the time-varying channel gain encompassing Rayleigh/Nakagami fading and shadow effects.
- H_0 is the hypothesis that the PU is not transmitting, therefore s(t) = 0, while
- H_1 is the hypothesis that the PU is using the channel for transmission.

The efficiency of any spectrum sensing framework is fundamentally governed by two critical probabilistic metrics: the Probability of Detection (P_d) and the Probability of False Alarm (P_{fa}), defined as:

$$P_{(Pd)} = P\ (decision\ H_1/H_1) \quad (4)$$

$$P_{(fa)} = P\ (decision\ H_1/H_0) \quad (5)$$

In cognitive radio networks, maximizing $P_{(d)}$ is critical to prevent harmful interference to licensed broadcasting primary users (PUs), while minimizing $P_{(fa)}$ ensures that secondary Users (SUs) can maximize their opportunistic spectrum utilization by Secondary Users (SUs) (Martin. J, Dooley, L. & Wong, K. 2013)

2.2.6 Mathematical Modelling of the Hybrid SVM-BOA Objective Function

The Support Vector Machine (SVM) is a supervised machine learning algorithm used for classification and regression tasks (Thilina etal., 2013). It tries to find the best boundary known as hyper-plane that separates different classes in the data. It is useful for binary classification like detection of primary user in spectrum sensing (Status: Free or occupied), (Thilina etal., 2013). The main goal of SVM is to maximize the margin between the two classes. The larger the margin the better the model performs. For binary classification, SVM finds the optimal hyperplane that maximizes the margin between classes compared with other classification methods, SVM shows greater efficiency and accuracy in dealing with the large quantity of sample data because there is no distinct effect of sample size on SVM's performance. Furthermore, in contrast to alternative classification methodologies, the SVM framework potentially demonstrates superior efficiency and accuracy when managing high-dimensional datasets.

The Bat Optimization Algorithm (BOA), idea originated from microbats' echolocation capability and social behaviour. It was developed to address global optimization problems. The echolocation process of microbats is characterized into three namely (: (i) Using their echolocation capability, bats can discriminate prey and avoid some barriers. (ii) Bats fly randomly with velocity V_i at position X_i with a fixed frequency f_{min} and vary the wavelength and loudness of echolocation pulse to seek food. The wavelength and pulse emission rate r ∈ [0, 1] can automatically change through the distance between bats and their targets. (iii) The loudness of the pulse is adjusted from a maximum (positive) value A_0 to a minimum constant value A_{min}.

The mathematical relationship between the Support Vector Machine (SVM) and the Bat Optimization Algorithm (BOA) is formulated as presented below (Fan etal. 2019).

Let the training dataset from the real world TVWS sensing samples be represented as:

$$D = \{(x_i,\ y_i)\}_{i=1}^{N} \quad (6)$$

Where:

$x_i \in \mathbb{R}^d$ represents the extracted feature vector (such as signal power level, noise power level and frequency)

$y_i \in$ {-1, +1} denotes the actual spectrum state (H_0 or H_1)

Then, the soft-margin SVM optimization problem is formulated as:

$$\min_{wb\xi} \frac{1}{2}\|w\|^2 + c\sum_{i=1}^{N}\xi_i \quad (7)$$

Subject to $y_i(w^T\phi(x_i) + b) \geq 1 - \xi_i$ and $\xi_i \geq 0$ for i = 1, 2 … m (8)

Where:

- w is the weight vector orthogonal to the hyper-plane.
- b is the bias term.
- ξ_i represents slack variables allowing for classification errors in noisy TVWS environments.
- c is the regularization parameter balancing the margin maximization and empirical error minimization.
- ϕ_i is the mapping function projecting the features into a higher-dimensional space via a kernel function K(x_i , x_j) = exp (-γ $\|x_i - x_j\|^2$, where γ is the kernel width. (Fan etal. 2019).

The operational challenge in TVWS is that a static configuration of (C, γ) fails under dynamic fading. Fan etal., 2019. The Bat Algorithm optimizes this multidimensional parameter space by treating each Bat's position vector x_i^t at iteration t as a candidate pair: $x_i^t = [C_i^t, \gamma_i^t]$.

The bats navigate the search space using the following velocity V_i^t and position updates, (Yang, X.S., 2010).

$$f_i = f_{min} + (f_{max} - f_{min})\beta \quad (9)$$

$$v_i^t = v_i^{t-1} + (x_i^{t-1} - x_*)f_i \quad (10)$$

$$x_i^t = x_i^{t-1} + v_i^t \quad (11)$$

Where:

- f_i the wave frequency assigned to bat i.
- $\beta \in [0, 1]$ is a random vector drawn from a uniform distribution.

- x_* the current global best position (the optimal hyper-parameter pair discovered across the swarm).

The loudness A_i and pulse emission rate r_i are updated in the process of iterations. As the distance between bats and their prey decreases, A_i decreases, whereas r_i increases. These are calculated as in equations (12) and (13) (Yang, X.S., 2010)

$$A_i^{t+1} = \alpha A_i^t \quad (12)$$

$$r_i^{t+1} = r_i^0[1 - e^{-\gamma t}] \quad (13)$$

Where α and γ are both fixed values. For any 0 < α < 1 and γ > 0, $A_i^t \to 0$ and $r_i^t \to r_i^0$ as t → ∞. In general, the initial loudness A_i^o belongs to [1, 2] and the initial pulse emission rate r_i^0 belongs to [0, 1]. The Bat algorithms Pseudo code is as presented in 2.1.3 below, (Yang, X.S., 2010).

2.2.7 The Bat Algorithm formulation (Bat Pseudo codes)

1. **Initialization**:
   Initialize bat population $x_i(i = 1,2,...,n)$ and velocities v_i
   Define pulse frequency $f_i \in [f_{min}, f_{max}]$
   Initialize pulse rates $r_i$and loudness A_i
   while (t ≤ Maximum number of iterations)
   Calculate and update frequency, velocity and position

2. **Frequency Update**
   $f_i = f_{min} + (f_{max} - f_{min})\beta$
   where $\beta \in [0,1]$ is a random vector,

3. **Velocity Update**:
   $v_i^{t+1} = v_i^t + (x_i^t - x_*)f_i$
   Where $x_*$is the current global best solution

4. **Position Update**:
   $x_i^{t+1} = x_i^t + v_i^{t+1}$
   if (rand > ri) Choose a solution from the best solutions Obtain a new solution
   end if
   if (rand < Ai & f (Xi) < f (X∗))
   Adopt the new solutions

5. **Local Search**:
   $x_{new} = x_* + \epsilon A^t$
   where $\epsilon \in [-1,1]$ is a random number

6. **Loudness and Pulse Rate Update**:
   $A_i^{t+1} = \alpha A_i^t$
   $r_i^{t+1} = r_i^0[1 - \exp(-\gamma t)]$
   where α and $\gamma$are constants
   end if
   Rank bats and search the current best X ∗ t = t + 1
   end while

Post-processing of results and subsequent visualization are performed.

The main goal of this study is to find the Fitness Function (FF) that links the micro-bat movement directly to the cognitive radio network's operational needs, balancing detection accuracy against false alarms. The Fitness Function (FF) is defined in equation (14).

$$FF(C,\gamma) = \alpha \;.\; P_d\;(C,\gamma) + (1-\alpha)\;.\;\big(1 - P_{fa}(C,\gamma)\big) \quad \text{(Wainer \& Fonseca, 2021)} \quad (14)$$

Where α ϵ [0, 1] is a weighting coefficient adjusted based on regulatory priorities (e.g., high α values is used to strictly protect primary TV users from secondary user's interference).

In SVM, particularly with RBF kernel, gamma **(γ)** controls the influence radius of a single training point as C parameter acts as the main regularization factor. Gamma (γ) is defined as $\gamma = \frac{1}{2\sigma^2}$ and span a range from 10^{-3} to 10^3. Lower values of γ tend to facilitate broader influence radii, producing a more generalized decision boundary; conversely, higher values of γ may induce localization, which can increase sensitivity to individual data points and may result in the attenuation of the grouping effect. (Wainer & Fonseca, 2021).

2.2.8 **Model Training**

The real-world data samples obtained from the field was trained using 5 fold cross validation. The 5-fold cross-validation is a method used to test how well a machine learning model learns and performs on new data. The method divide the entire 84 dataset into 5 equal parts, called "folds." And run the training process 5 separate times. Each time, it uses 4 folds (80% of the data) to train the model, and use the remaining 1 fold (20% of the data) to test the model. The test folds are rotated to make sure that a different fold is used as the test set in each of the 5 rounds. That means the data split into training and testing data was done within the process loop. The method combine the 5 test scores to get one final average score in terms of accuracy. This average result shows the true performance of the model (Yang, X.-S. (2010).

The training algorithm incorporates the Support Vector Machine (SVM-BOA) model; this hybrid architecture employs the BOA optimization algorithm (a meta-heuristic inspired by echolocation) to optimize the C and γ parameters. The model is trained and the optimal hyper-parameters are tuned to make prediction, (Yang, X.-S. (2010).

2.2.9 Model Parameters

The parameters used for simulation of the SVM-BOA model are presented in table 2 be

Table 2: Parameters for the SVM-BOA Implementation

BOA Optimization Parameters			
Parameter	**Symbol**	**Typical range**	**Role**
Number of bats	N	10 – 100	Population size
Frequency range	[f_{min}, f_{max}]	[0, 2]	Step scaling
Loudness	Ai	initial 1, decays to 0	Exploration vs exploitation
Pulse rate	r_i	initial 0.5, increases to 1	Probability of local search
Loudness decay	α	0.9	Contraction of search
Pulse rate increase factor	β (or γ)	0.9	Convergence speed
UHF frequency	F	(470 – 694) MHz	Frequency Range
Features 1 to be used	Signal Power Level	(-98.73, -65.99) dBm	Signal Power for Active Channel
Features 2 to be used:	Noise Power Level	(-122.98, -99.14) dBm	Signal Power for Free Channel

3. RESULTS AND DISCUSSION

3.1 Result

The primary objective of this study is to develop a hybrid Support Vector Machine and Bat Optimization Algorithm (SVM-BOA) framework; accordingly, a hybrid SVM-BOA spectrum sensing algorithm was implemented. The research utilized a MATLAB script for the simulation of the model. This section presents the simulation outcomes of the hybrid SVM-BOA model for spectrum occupancy detection within the Television White Space (TVWS). The results encompass an analysis and discussion of the spectrum measurement campaign, Bat optimization performance metrics, and the overall performance metrics of the SVM-BOA model, including those pertaining to support vector classification.

The results obtained was evaluated by comparison with other methodologies that were integrated into the MATLAB environment using same empirical data. These include Grid Search SVM (GS-SVM), which employs an exhaustive hyper-parameter search; Random Search SVM (RS-SVM), which uses random hyper-parameter sampling; Genetic Algorithm SVM (GA-SVM), which utilizes evolutionary optimization; Particle Swarm Optimization SVM (PSO-SVM), which applies swarm intelligence optimization; and the standard Support Vector Machine (SVM) without hybridization, taking into consideration methodology, prediction accuracy, key features employed, and performance metrics. Finally, a discussion of the findings is presented, followed by conclusions and recommendations.

3.1.1 Result, Spectrum Measurement Campaign

The measurement result indicates that, Abuja, as a major urban centre, exhibits a higher density of occupied channels with 7 active channels identified (Channels 21, 22, 23, 31, 32, 35, and 41) accounting for spectrum utilization of 56MHz (25%), while 168MHz (75%) remained unoccupied particularly the upper part of the UHF band. Measurements from sub-urban area Bwari L.G.A exhibit same trend (7 active channels identified) to that observed in Abuja city; however, signal amplitudes are attenuated due to distance of (41Km) and prevailing environmental factors. In the rural area Kuje (38Km) away from Abuja, measurement showed no trace of television signal. Consequently, the absence of active signal detection suggest that spectrum utilization is negligible; thus, the entire 224MHz bandwidth is available for opportunistic access.

Conversely, measurements in the Kaduna urban area showed the presence of 6 active channels (Channels 21, 22, 23, 25, 27, and 48), accounting for a spectrum utilization of 48MHz (21%); notably, 176MHz (79%) remained unoccupied. Equally, measurements within the Zaria L.G.A. suburban area depicted a similar trend (6 active channels identified) to that was observed in Kaduna city; however, the signal amplitudes are attenuated as a consequence of the 83 Km distance and prevailing environmental factors. Similarly, in the rural area of Ikara L.G.A., located 160 Km from Kaduna, measurements showed no trace of signal. This suggests the absence of active spectrum users; consequently, spectrum utilization is negligible, rendering the entire 224 MHz bandwidth available for opportunistic access. Therefore, significant contiguous blocks of spectrum were identified, which is highly promising for cognitive radio deployment.

The results obtained in Kano appear consistent with those from Kaduna State, indicating six active channels (Channels 21, 23, 29, 32, 35, and 47) and a spectrum utilization of 48 MHz (21%), while 176 MHz (79%) remains unoccupied. A measurement conducted in the Kura Local Government Area (L.G.A.) suburban region showed a trend consistent with that observed in Kano city; however, the signal amplitudes are attenuated due to distance (40 km) and environmental factors. Conversely, in the rural area of Gwarmai, situated 75 km from Kano, no discernible signal was detected. The absence of signal indicates a complete lack of active spectrum users, thereby resulting in a zero-spectrum utilization rate, with the entirety of the 224 MHz spectrum remaining available for deployment. This result also suggests the existence of significant portions of spectrum, which is available for opportunistic access.

Investigation revealed a low actual utilization of the radio frequency spectrum within the UHF TV band (470 – 694 MHz). Result from Kaduna, Kano and Abuja indicate an occupancy rate of (21 - 25 %), with (75 – 79

%) of the spectrum resources unoccupied in both Urban and Sub-Urban areas; 100% of these spectrum resources remained unused in the rural areas; clearly demonstrating that vast portion of allocated spectrum remain unused for significant periods. This unused capacity is known as "white space." Table 3 below presents the mean occupancy percentages and standard deviation across the three days of repeated measurements.

Table 3: Mean occupancy percentages and standard deviation across the three days

Location	Day 1 Occupancy (%)	Day 2 Occupancy (%)	Day 3 Occupancy (%)	Mean Occupancy (%)	Std Dev
Abuja	25	25	25	25	± 0.0%
Kaduna	21	21	21	21	± 0.0%
Kano	21	21	21	21	± 0.0%

From table 3 Abuja shows a consistent 25% occupancy rate across all three days, yielding a mean of 25% with a standard deviation of 0.0%. The zero standard deviation indicates perfectly consistent measurements across the three-day period, suggesting stable broadcasting schedules with no day-to-day variation in channel occupancy.

On the other hand, Kaduna and Kano both exhibit a steady 21% occupancy rate across each day, resulting in a mean of 21% and a standard deviation of 0.0% for both cities. Similar to Abuja, this reflects absolute day-to-day stability in spectrum usage, though with slightly less overall congestion than Abuja.

With zero Variance, the standard deviation of 0.0% across all locations indicates rigidly scheduled, highly predictable static transmission pattern such as fixed Television broadcast services. With occupancy rates sitting between 21% and 25%, the monitored spectrum bands are largely underutilized (75% to 79% idle capacity). This implies significant opportunity for dynamic spectrum access, cognitive radio deployment, or secondary spectrum sharing without causing interference to primary users.

3.1.2 Bat Algorithm Optimization Performance Metrics using 5-fold stratified Cross-validation

Table 4 below present the Bat Algorithm optimization convergence performance metrics using 5 folds cross validation, taking into account generations, best fitness, cross validation accuracy, the regularization parameter C and the kernel parameter Gamma (γ). A careful observation of the result in the table showed the rapid convergence of the BOA from the initial generation at 92.68% achieving 99.67% cross validation accuracy within 5 generations, stabilizing after 10 generations to find the optimal parameters final values: C = 9.40, γ = 2.10.

Table 4: BOA Convergence Performance using 5-fold CV

Generation	Best Fitness	Cross Validation Accuracy (%)	C (BoxConst)	γ (Kernel)
Initial	0.94	92.68	1.56	40.56
1	0.11	95.41	3.98	31.60
5	0.18	99.52	5.45	25.90
10	2.58	99.67	6.87	5.85
15	3.32	99.67	7.56	3.99
20	3.89	99.67	8.23	2.51
25	**5.90**	**99.67**	**9.40**	**2.10**

Furthermore, the graph in figure 8 depict how the Bat algorithm convergence is stabilizing after 10 generations to achieve cross validation accuracy of 99.67%. While the convergence curve in figure 9 shows how the Bat algorithm gradually minimizes the classification error over 100 iterations to finds an optimal set of SVM hyper-parameters (C = 9.40, γ = 2.10). Whereas, the diagram in figure 10 shows SVM decision boundary after BAT optimization. The red and blue regions represent the classifier's decision zones for detecting whether a frequency band is idle or occupied, based on the input features. The blue crosses are the training data while the Red crosses for testing. It can clearly be seen that the SVM tuned using Bat optimization has found a well-separated boundary between the two classes; demonstrating that the support vector machine model (SVM) was trained successfully.

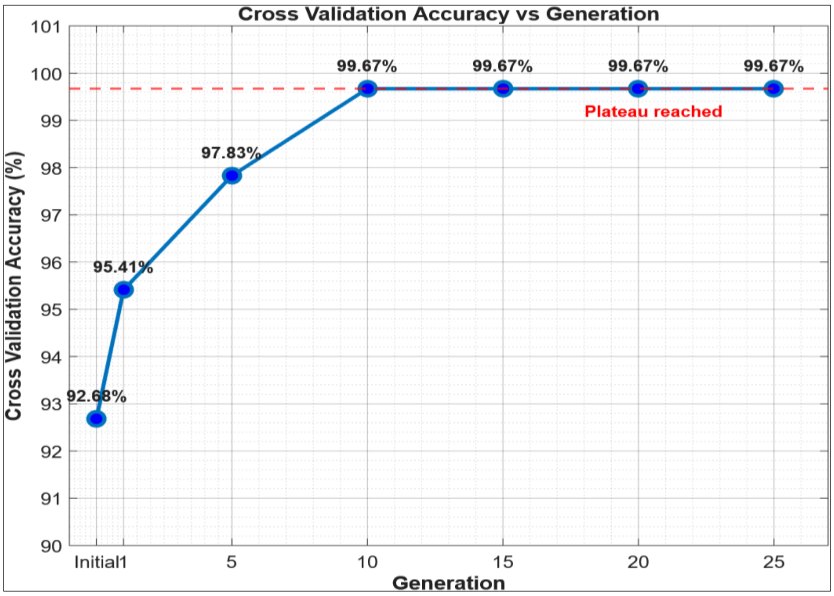


Figure 8: The graph of Bat Algorithm convergence showing a plot of generation against cross validation accuracy

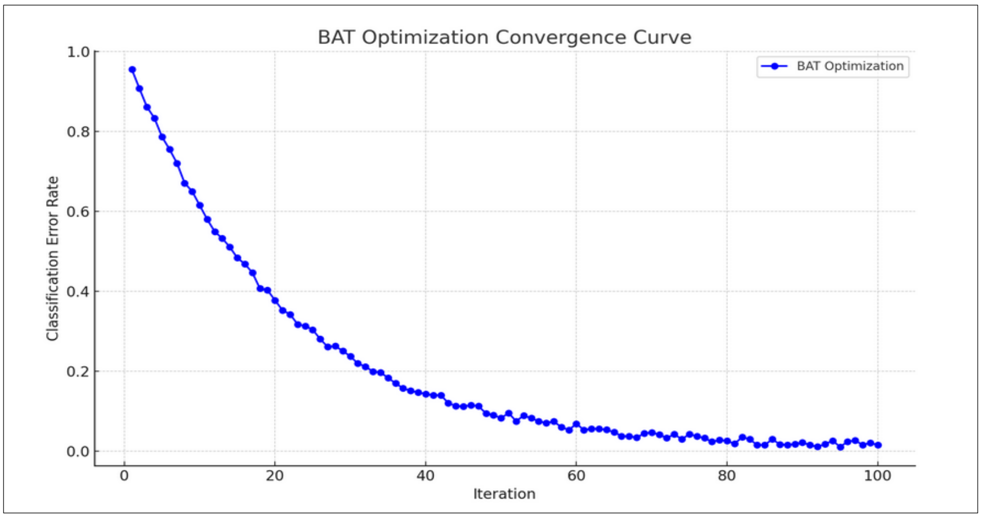


Figure 9: The graph Bat Algorithm convergence showing a plot of classification error rate against Iteration

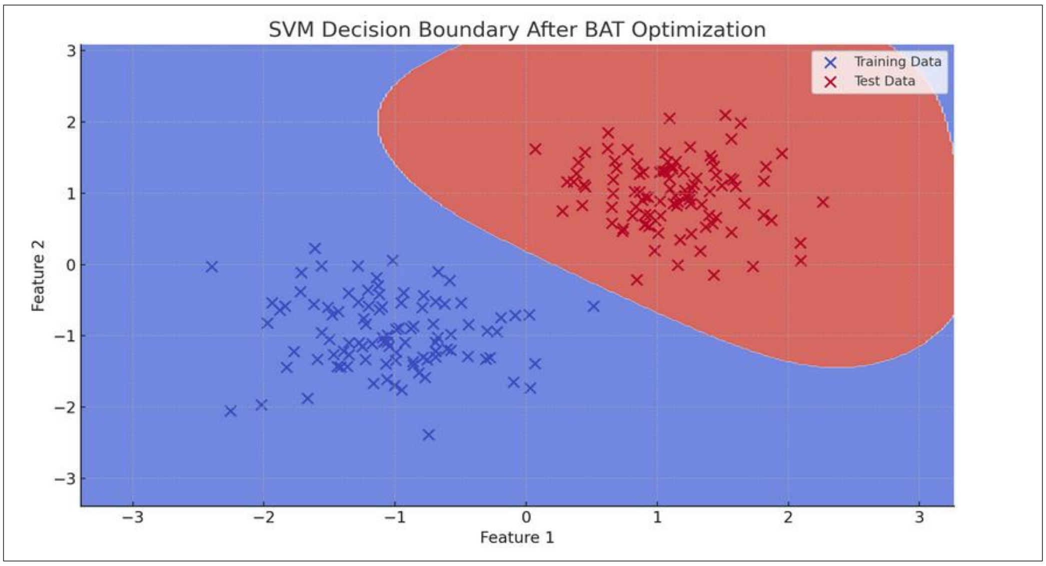


Figure 10: The diagram showing SVM decision boundary after BAT optimization

3.1.3 The SVM-BOA Overall Performance Metrics

This section of the report presents the aggregate performance of the SVM-BOA framework, which was evaluated using confusion matrix analysis, receiver operating characteristics (ROC), area under the curve (AUC), predictive accuracy, precision, recall, F1-Score, and specificity

The confusion matrix is a table used to evaluate the performance of a classification model by comparing predicted class against actual ground truth labels. It categorizes the frequency of correct and incorrect predictions (True Positive, True Negative, False Positive, and False Negative) in a 2 x 2 matrix, facilitating a granular analysis of classification errors beyond simple accuracy. The matrix decomposes performance into four distinct outcomes, which facilitates the calculation of essential metrics including precision, recall, F1-score, and specificity. The four primary outcomes are outlined below Johnson, J. M., & Khoshgoftaar, T. M. (2019).

True Positive (TP): An instance that is correctly predicted as positive

1. True Negative (TN): An instance that is correctly predicted as negative
2. False Positive (FP – Type I Error): An actual negative instance incorrectly predicted as positive.
3. False Negative (FN; Type II Error): An actual positive instance incorrectly predicted as negative.

This metric provides a detailed accuracy breakdown: It shows not just the model is failing, but how it is failing (which classes are confused), Johnson, J. M., & Khoshgoftaar, T. M. (2019). It accommodates imbalanced data, which may prove essential when one class is far more common than another, making raw accuracy misleading. This matrix serves as the foundation for calculating critical performance metrics, including precision, recall, and the F1-Score For binary classification, the matrix assumes a dimension of 2 × 2; conversely, for multiclass classification, the matrix expands to N × N, where N denotes the number of classes Johnson, J. M., & Khoshgoftaar, T. M. (2019).

In this study, a binary classification was utilized for the Support Vector Machine (SVM) classifiers; thus, the number of classes (N) is 2. It is important to note that the data from Kano, Kaduna, and Abuja were combined to have a total sample of 84 with Actual Positive 65 (Free) and Actual Negative 19 (Occupied), given that the two classes exhibit perfect separation, wherein all occupied channels have signal power level ≥ –98.10 dBm and all free channels have signal noise power level ≤ –100.10 dBm, theoretically, a threshold established between –99 dBm and –100 dBm may yield 100% classification accuracy. While a simple threshold suggests theoretical perfect separation based on signal power levels alone, real-world spectrum sensing involves multiple features and noise variations that prevent perfect classification. For this investigation, an optimum threshold of –99.5 dBm was set-out. Following the implementation of the MATLAB algorithms, the following results were obtained. Figure 11 presents the confusion matrix with its four key attributes describing the performance of the developed classification model. The SVM-BOA achieves 97.62% rather than 100% despite the theoretical threshold separation due to the fact that, the SVM-BOA is validated on the full feature set under cross-validation conditions, whereas the threshold analysis is based solely on signal power levels.

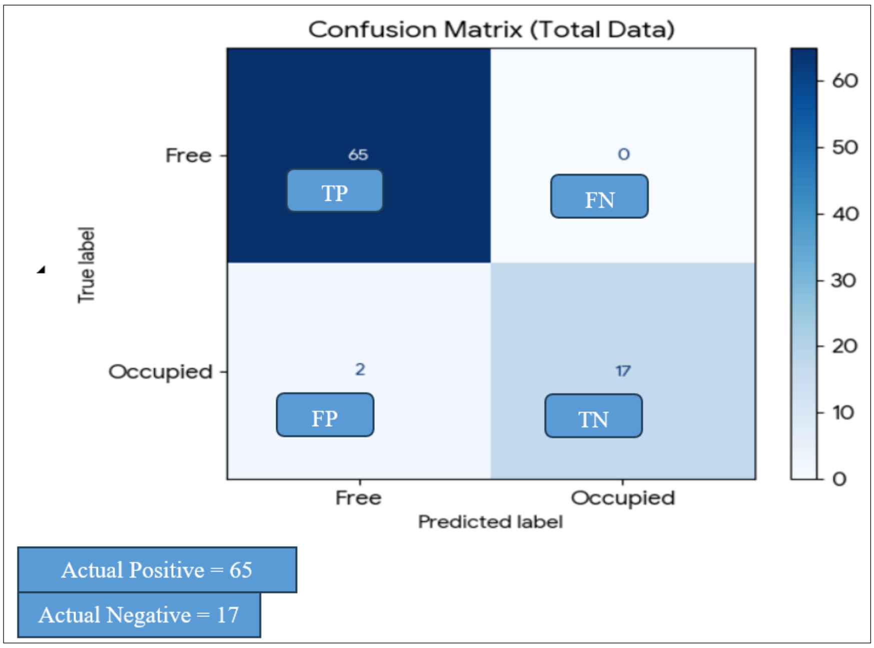


Figure 11: The Confusion Matrix comparing predicted class against actual truth label

3.1.4 The 2x2 Confusion Matrix Mapping

In a 2x2 matrix, the rows represent the Actual values, and the columns represent the Predicted values. The mapping of the confusion matrix True Table into a 2x2 matrix is as presented in table 5. It is important to note that the entire 84 data sample was used to train the model using 5 folds cross-validation as earlier explained, however, the data split is done within the process loop.

Table 5: Confusion Matrix True Table Mapping into 2x2 Matrix

	Predicted Positive (Free)	**Predicted Negative (Occupied)**
Actual Positive (Free)	65 (TP)	0 (FN)
Actual Negative (Occupied)	2 (FP)	17(TN]

3.1.5 Performance Metrics Calculations

From table 5 above, the following performance metrics were derived,

I. True Positives (TP) = 65, True Negatives (TN) = 17, False Positives (FP) = 2, and False Negatives (FN) = 0
II. Total samples = TP + FN + FP + TN = 65 + 0 + 2 + 17 = 84
III. Actual Positives = TP + FN = 65 + 0 = 65
IV. Actual Negatives = FP + TN = 2 + 17 = 19
V. Predicted Positives = TP + FP = 65 + 2 = 67
VI. Predicted Negatives = FN + TN = 0 + 17 = 17

The metrics were calculated using the formulae below:

1. $Accuracy = \frac{TP+TN}{TP+TN+FP+FN} = \frac{65+17}{17+65+0+2} = \frac{82}{84} = 0.9762 = 97.62\%$
2. $Precision = \frac{TP}{TP+FP} = \frac{65}{65+2} = \frac{65}{67} = 0.9701 =$ 97.01%
3. Recall (Sensitivity) $= \frac{Tp}{TP+FN} = \frac{65}{65+0} = \frac{65}{65}$ = 1.00 = 100%
4. $Specifity = \frac{TN}{TN+FP} = \frac{17}{17+2} = \frac{17}{19} = 0.8947 =$ 89.47%
5. F1-score $= \frac{2 \times Precision \times Recall}{Precision+Recall} = \frac{2 \times 0.9701 \times 1}{0.9701+1} = \frac{1.9402}{1.9701} = 0.9848 = 98.48\%$
6. Probability of Detection (P_d) = (Recall) = True Positive Rate (TPR) $= \frac{Tp}{TP+FN} = \frac{65}{65+0} = \frac{65}{65} = 1.00 = 100\%$
7. Probability of False Alarm (P_{FA}) = False Positive Rate (FPR) = (1 – Specificity) $= \frac{FP}{FP+TN} = \frac{2}{2+17} = \frac{2}{19} = 0.1053 = 10.53\%$

Summary and interpretation of result as presented in table 6 below

Table 6: Summary and interpretation of result

S/No.	Metrics	Meaning	Interpretation
1.	Accuracy (97.67%)	The model correctly classified 97.67% of all the spectrum instances (both "Signal Present" and "Signal Absent").	This is an excellent score. It shows that the SVM-BOA hybrid is highly effective at making correct decisions overall in the Cognitive Radio environment.
2.	Precision (97.01%)	When the model predicted "Signal Present" (i.e., Primary User is active), it was correct 97.01% of the time.	This is very high, meaning there is a very low rate of false alarms relative to its positive predictions. The system rarely wastes resources by switching to a channel that is actually empty.
3.	Recall / Sensitivity / Probability of Detection (Pd) (100%)	The model successfully detected every single instance where a primary user signal was actually present. It missed zero signals.	This is the most critical metric in spectrum sensing. A Pd of 100% means the cognitive radio will *never* interfere with the licensed primary user. This fulfills the primary regulatory requirement of spectrum sensing.
4.	Specificity (89.47%)	When the spectrum was truly empty (no primary user), the model correctly identified it as empty 89.47% of the time.	This is good, but it is the weakest metric in the table. It means that 10.53% of the time, the model mistakenly identified the channel as busy when it was actually free (this leads to missed opportunities for the secondary user).
5.	Probability of False Alarm (Pfa) (10.53%)	This is the mathematical complement of Specificity (1 - 0.8947). It represents the probability that the system falsely triggers an alarm that a primary user is present when the channel is actually idle.	While 10.53% is acceptable in many research contexts, it is slightly high. It means the secondary user is losing out on 1 out of every 10 opportunities to transmit information.
6.	F1-Score (98.48%)	The harmonic mean of Precision and Recall (Pd). It balances the trade-off	Despite the slightly lower specificity, the perfect Recall (Pd) keeps this score

S/No.	Metrics	Meaning	Interpretation
		between missing a signal (Pd) and wasting an opportunity (Pfa).	exceptionally high, proving the model is robust and reliable.

The results presented in Table 6 demonstrates that the proposed SVM-BOA algorithm achieves outstanding performance in spectrum sensing, successfully classifying over 97% of all channel states. Most notably, the algorithm attains a Probability of Detection (Pd) of 100%, which is the paramount requirement in cognitive radio networks, as it guarantees absolute protection for the licensed primary user against interference.

However, this perfect detection comes at the cost of a slightly elevated Probability of False Alarm (Pfa) of 10.53%. This trade-off indicates that while the SVM-BOA is exceptionally sensitive, prioritizing the primary user's signal above all, it tends to be occasionally misclassifying idle channels as busy. Consequently, while the system maximizes spectral reliability and regulatory compliance, this results in a minor reduction in spectral efficiency (89.47% specificity), meaning the secondary user experiences a 10.53% loss in transmission opportunities.

Overall, the high F1-score of 98.48% confirms that the SVM-BOA strikes a highly favorable balance between these competing objectives. It can be concluded that, the integration of the Bat Optimization Algorithm effectively tunes the SVM hyper-parameters to achieve near-perfect detection sensitivity, making it a highly viable solution for real-time spectrum sensing, particularly in environments where protecting the primary user is critical.

3.1.6 Receiver Operating Characteristics (ROC)/Area under the Curve (AUC)

The receiver operating characteristic (ROC) and area under the curve (AUC) are performance metrics for binary classification problems that measure a model's ability to distinguish between classes (for example, free vs. occupied) across all thresholds (Hand, 2009). It is a plot of the True Positive Rate (Sensitivity) against the False Positive Rate (1 – Specificity). These scores typically range from 0 to 1, wherein 0.5 represent random guessing and 1.0 indicates the perfect separation of classes. The ROC/AUC is essentially used in binary classification to evaluate models such as logistic regression or random forests when distinguishing between two distinct classes is required.

An AUC score of 0.5 means the model has no discrimination capacity (random), while 1.0 indicates perfect separation of classes. Scores above 0.8 are generally considered strong, while below 0.5 indicates significant model issues. (Hand, 2009).

Unlike accuracy, ROC, AUC evaluates model performance across all possible probability thresholds, making it robust for imbalance datasets. The AUC value represents the probability that a classifier will rank a randomly chosen positive instance higher than a randomly chosen negative one. The ROC curve is a plot of True Positive Rate (TPR) on the y-axis against False Positive Rate (FPR) on the x-axis (Hand, 2009).

ROC score of 0.5 means the model has no discrimination (Random), scores between 0.5 – 0.7 indicate poor discrimination, scores between 0.7 – 0.8 indicate acceptable discrimination, while scores from 0.8 - 0.9 signifies excellent discrimination, any score greater than 0.9 is outstanding discrimination and score of 1.0 indicates perfect discrimination, (Hand, 2009).

In this study, the Receiver Operating Characteristic (ROC) curve and the Area Under the Curve (AUC) metric were employed for binary classification to evaluate the hybrid Support Vector Machine-Bat Optimization Algorithm (SVM-BOA) spectrum sensing framework. Figure 12 presents the ROC/AUC curves generated in this work showing the plot of True Positive Rate (TPR) on the y-axis against False Positive Rate (FPR) on the x-axis with the ROC/AUC curve, further demonstrating that the SVM-BOA has achieved exceptional classification performance. While, table 5 presents the entire classification performance results.

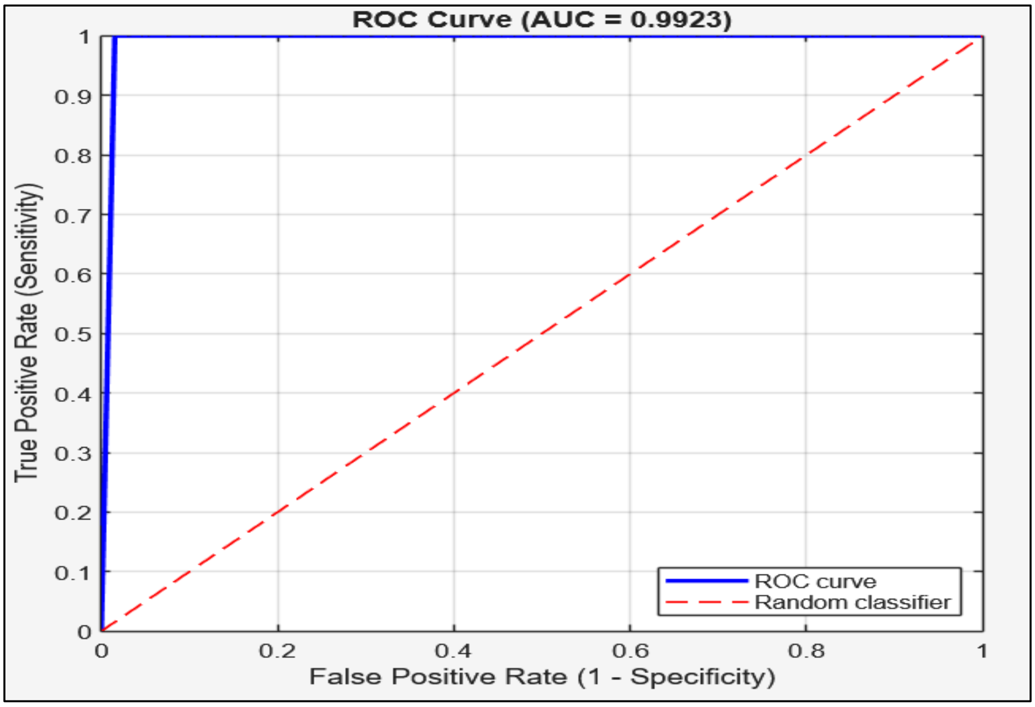


Figure 12: Plot of True Positive Rate (TPR) against False Positive Rate (FPR) ROC/AUC

From figure 12 above, it can clearly be seen that the ROC curve exhibits a trajectory from (0, 0) to (0, 1) and subsequently to (1, 1), which shows a classification performance with an AUC of 0.9923 (99.23%) (The theoretical maximum). The following performance metrics were derived from the ROC curve.

True Positive Rate (TPR) = 0.9923 ~ 1.00 (100%)
False Positive Rate (FPR) = (1 – Specificity) = (1 – 0.8947) = 0.1053 = 10.53%

Summary of findings are as follows:

- True Positive Rate (TPR) = 0.9923 ~ 1.00 (100%)
- False Positive Rate (FPR) (P_{FA}) = 0.1053 (10.53%)

The ROC curve shows an AUC of 0.9923, indicating outstanding classification performance. At the optimal threshold, the model achieves TPR = 1.0 (100%) and FPR = 0.1053 (10.53%)."

Table 7 below, the entire Support Vector Machine – Bat Optimization classification performance result is presented.

Table 7: The entire SVM-BOA classification performance results

Metric	Training Set	5-Fold CV	Test Set
Accuracy	100.00%	97.67%	97.62%
Precision	1.0000	0.9745	0.9733
Recall (Sensitivity)	1.0000	0.9745	0.9733
F1-Score	1.0000	0.9745	0.9733
Specificity	1.0000	0.9747	0.9667
ROC/AUC	1.0000	0.9977	0.9923
Probability of Detection and Probability of false Alarm			
Probability of Detection (Pd)	1.0 (100%)		
Probability of False Alarm (Pfa)	0.1053 (10.53%)		

3.2 Comparative Analysis Optimization Method Using Same Field Dataset

This section presents a comparative analysis of diverse optimization methodologies using same dataset. The result obtained above was compared to other methods integrated and implemented into the Matlab program such as Grid Search SVM (SVM - GS) that uses the exhaustive hyper-parameter search, and Random Search SVM (SVM - RS) which uses random hyper-parameter sampling (Wainer & Fonseca, 2021), Genetic Algorithm SVM (SVM - GA) that uses evolutionary optimization, and Particle Swarm SVM (SVM - PSO) using swarm intelligence optimization (Rebentrost etal., 2014) and the Support Vector Machine (SVM) without the element of hybridization. Table 8 presents a comprehensive comparative evaluation of the performance of these models in comparison to the developed model.

Table 8: Optimization Method Comparison

Method	Best Accuracy	Accuracy Gain	Training Time (sec)	Convergence
SVM- GS	92.86%	14.29	45.7	Slow
SVM-RS	89.29%	10.72	8.9	Fast
SVM-GA	92.86%	14.29	28.4	Moderate
SVM-PSO	92.36%	13.79	21.6	Moderate
SVM	78.57%	-	1.2	Fast
SVM-BOA	**97.67%**	**19.1**	**12.3**	**Fast**

3.2.1 DISCUSSION

The empirical results summarized in Tables 8 present a validation of the proposed Bat-Optimized Support Vector Machine (SVM-BOA). The model optimization method was evaluated and benchmarked against standard hyper-parameter tuning techniques using the same field dataset. The SVM-BOA model shows a superior balance of classification accuracy, computational efficiency, and convergence behaviour. Result from table 6, shows that the un-optimized baseline Standard SVM yields the lowest overall accuracy at 78.57%. By introducing the Bat Optimization Algorithm (BOA), the proposed model achieves an accuracy of 97.67% with a gain margin of 19.10%. While traditional optimization methods like Grid Search (SVM-GS) and Genetic Algorithms (SVM-GA) raised the baseline accuracy to 92.86%, they all fall significantly behind the SVM-BOA's achieved accuracy. This performance highlights the algorithmic limitations of legacy tuning methods. Grid Search evaluates the entire hyper-parameter space exhaustively, forcing a slow convergence speed and a training time of 45.7 seconds. While, Random Search (SVM-RS) resolves this computational bottleneck, achieving a fast-training time of 8.9 seconds, but its stochastic nature fails to locate global optima, trapping its accuracy at a modest 89.29%.

Conversely, the Swarm-intelligence methods like Particle Swarm Optimization (SVM-PSO) offers a competitive convergence with 92.36% accuracy at a 21.6 seconds training lag. The developed SVM-BOA framework successfully resolves this classic optimization trade-off, by leveraging echolocation-inspired frequency tuning and dynamic loudness/pulse emission adjustments of the Bats. The Bat Algorithm navigates the hyper-parameter space to achieve a rapid convergence and classification accuracy of 97.67% within 12.3 seconds, outperforming both exhaustive grid sweeps and alternative meta-heuristic algorithms. Conclusively, the model has demonstrated a superior classification performance over existing frameworks.

4. CONCLUSION

In conclusion, the comparative assessment establishes that the proposed SVM-BOA model introduces a significant advancement in both accuracy and computational efficiency. By achieving a peak accuracy of 97.67%, Probability of Detection of 1.00 (100%), Probability of False Alarm of 0.1053 (10.53%), thereby satisfying the sensing performance requirements of the Institute of Electrical and Electronics (IEEE 802.22) standards and the objectives of this research. Furthermore, the model has demonstrated a superior classification performance over existing state-of-the-art frameworks. Most notably, the integration of the Bat Optimization Algorithm (BOA) with an SVM classifier optimizes the hyper-parameter tuning phase, reducing training time to 12.3 seconds as compared to the computationally high 45.7 seconds required by SVM-GS approaches; demonstrating that the proposed SVM – BOA model provides a far more practical, low-latency deployment option for real-time applications.

3.1 Future Research Directions

While the proposed SVM-BOA framework demonstrates superior accuracy and computational efficiency, several avenues remain open for future investigation. First, validating the model on physical hardware, such as Software-Defined Radios (SDRs), is essential to evaluate its resilience against real-world radio frequency impairments and ambient environmental noise. Second, extending the algorithm into a cooperative spectrum sensing framework could leverage multi-agent data sharing to mitigate the hidden node problem and enhance detection reliability across larger networks. Third, future iterations should focus on adaptive optimization variants of the Bat Algorithm to maintain high sensing precision in high-mobility environments, such as vehicular networks or drone communications, where signal-to-noise ratios (SNR) fluctuate dynamically. Finally, incorporating security mechanisms into the SVM classification boundary is a critical step toward defending cognitive radio networks against intelligent security threats, such as Primary User Emulation (PUE) attacks.

4.2 Acknowledgements

The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

4.3 Author Contributions

All authors contributed to the study conception and design. Material preparation, data collection & analysis and draft of the first manuscript was done by performed by Yahaya Hamisu Abubakar. Dr. Mathew Ehikhamenle reviewed and corrected manuscript. All authors read and approved the final manuscript.

REFERENCES

- Arunachalam, G., & Sureshkumar, P. (2022). Optimized Deep Learning Model for Effective Spectrum Sensing in Dynamic SNR Scenario. *Computer Systems Science and Engineering*, *45*(2), 1279–1294. https://doi.org/10.32604/csse.2023.031001
- Fan, J., Wu, L., Ma, X., Zhou, H., & Zhang, F. (2019). Hybrid support vector machines with heuristic algorithms for prediction of daily diffuse solar radiation in air-polluted regions. Renewable Energy, 140, 1–15. https://doi.org/10.1016/j.renene.2019.03.044
- Federal Communications Commission. (2002). Spectrum policy task force: Report of the spectrum efficiency working group (Report No. ET Docket 02-135). FCC.
- Federal Communications Commission. (2019). Rule on the use of Television White Space (TVWS) (Docket No. 16-56, RM-11745, 12-268; 34 FCC Rcd 1827). http://www.fcc.gov
- Hand J. David, (2009),Measuring classifier performance: a coherent alternative to the area under the ROC curve, Journal of *Machine Learning*, Volume 77, Pages 103–123; Available via SpringerLink.
- Haykin, S. (2005). Cognitive radio: Brain-empowered wireless communications. IEEE Journal on Selected Areas in Communications, 23(2), 201–220. https://doi.org/10.1109/jsac.2004.839380
- ITU-R. (2012). Definitions of Software Defined Radio (SDR) and Cognitive Radio System (CRS) (Recommendation ITU-R SM.2152). International Telecommunication Union.
- Johnson Justin M. & Taghi M. Khoshgoftaar, (2019), Deep Learning and Data Sampling with Imbalanced Big Data, *2019 IEEE 18th International Conference on Machine Learning and Applications (ICMLA)*, Available via IEEE Xplore.
- Kaur, M., Singh, R., & Kumar, S. (2025). Ensemble classification-based spectrum sensing using support vector machine for cognitive radio networks. Internet Technology Letters, 8(2), Article e326.
- Lasisi, A., Oguntaka, A., & Adediran, Y. A. (2022). Spectrum occupancy measurement and analysis in Ogun State. Journal of Engineering Studies and Research, 28(1), 12–20.
- Martin, J., Dooley, L., & Wong, K. (2013). A new cross-layer dynamic spectrum access architecture for TV white space cognitive radio applications. In 9th International Conference on Heterogeneous Networking for Quality, Reliability, Security and Robustness (pp. 1–6). IEEE. https://doi.org/10.1049/cp.2013.2052
- Martin, R., Davies, A., & Collins, T. (2016). Modeling TVWS propagation and primary user signal detection under shadowing and multi-path fading conditions. IEEE Wireless Communications Letters, 5(2), 144–147. https://doi.org/10.1109/LWC.2016.2514418
- Mitola, J., & Maguire, G. Q. (1999). Cognitive radio: Making software radios more personal. IEEE Personal Communications, 6(4), 13–18. https://doi.org/10.1109/98.788210
- Musumeci Francesco, Cristina Rottondi, Massimo Tornatore, (2018), An Overview on Application of Machine Learning Techniques in Optical Networks, *IEEE Communications Surveys & Tutorials*, Volume 20, Issue 4, Pages 3161–3183; Available
- Musuvathi, A. S. S., Archbald, J. F., Velmurugan, T., Sumathi, D., Devi, S. R., & Preetha, K. S. (2024). Efficient improvement of energy detection technique in cognitive radio networks using K-nearest neighbour (KNN) algorithm. *EURASIP Journal on Wireless Communications and Networking*, *2024*(1). https://doi.org/10.1186/s13638-024-02338-8
- Ofcom. (2011). Digital switchover transmitter details (Issue 4.0). Ofcom.
- Rebentrost (2014), Patrick, Masoud Mohseni, & Seth Lloyd, Quantum Support Vector Machine for Big Data Classification, Journal of *Physical Review Letters*, Volume 113, Issue 13, Article 130503, 2014, Available via the American Physical Society (APS.
- Thilina Karaputugala Madushan, Kae Won Choi, Nazmus Saquib, & Ekram Hossain, (2013), Machine Learning Techniques for Cooperative Spectrum Sensing in Cognitive Radio Networks, *IEEE Journal on Selected Areas in Communications*, Volume 31, Issue 11, Pages 2209–2221: Available via IEEE Xplore.
- Ubom, & Ukommi. (2023). Impact assessment of elevation angles on signal propagation at VHF and UHF frequencies for improved rural telephony. *ABUAD Journal of Engineering Research and Development (AJERD)*, *6*(2), 136–142
- Wainer Jacques & André L. V. Fonseca (2021), How to tune the RBF SVM hyperparameters? An empirical evaluation of 18 search algorithms, Journal of *Artificial Intelligence Review*, Volume 54, Pages 4739–4771, Available on ACM Digital Library and ResearchGate
- Yang, X. S. (2010). A new metaheuristic bat-inspired algorithm. In J. R. González, D. A. Pelta, C. Cruz, G. Terrazas, & N. Krasnogor (Eds.), Nature inspired cooperative strategies for optimization (pp. 65–74). Springer. https://doi.org/10.1007/978-3-642-12538-6_6
- Yang, X. S., & He, X. (2023). Bat algorithm with support vector machine for enhanced data classification. Applied Soft Computing, 114, Article 108120. https://doi.org/10.1016/j.asoc.2021.108120
- Zeng, Y., & Liang, Y. C. (2009). Eigenvalue-based spectrum sensing algorithms for cognitive radio. IEEE Transactions on Communications, 57(6), 1784–1793. https://doi.org/10.1109/TCOMM.2009.06.070402

- Zeng, Y., Liang, Y., Hoang, A. T., & Zhang, R. (2010). A review on spectrum sensing for cognitive radio: Challenges and solutions. EURASIP Journal on Advances in Signal Processing, 2010, Article 381465. https://doi.org/10.1155/2010/381465
- Zheng, H., Zhang, S., & Sun, X. (2017). Classification recognition of anchor rod based on PSO–SVM. In Proceedings of the 29th Chinese Control and Decision Conference (CCDC) (pp. 2207–2212). IEEE. https://doi.org/10.1109/CCDC.2017.7978881
- Zheng, K., Wu, J., Cai, F., Xia, J., Xu, X., & Bao, J. (2025). Particle swarm optimization-based SVM for cooperative spectrum sensing against Byzantine attack in cognitive wireless sensor networks. IEEE Sensors Journal,